



NOAA
FISHERIES
SEFSC

QA/QC and Management of Data Collected during Fishery-Independent Surveys of the Southeast Fishery Science Center

MSRA Program Review

June 5, 2013

Presentation Layout

This presentation consists of two main sections:

QA/QC of fishery-independent survey data
and

Management of fishery-independent survey data

Presentation Layout cont.

For discussion of **QA/QC** of the numerous survey types employed by the SEFSC, the FI surveys to be discussed today are divided into the following groups:

Trawling Surveys

Plankton Surveys

Longline Surveys

Video, Trap, and Vertical Line Surveys

Visual Surveys, Caribbean Surveys, and Other Surveys

Presentation Layout cont.



NOAA
FISHERIES
SEFSC

For discussion of **QA/QC**, the following items are covered for our main large-scale surveys:

- Summary
- QA/QC of data

QA/QC of Fishery-Independent Survey Data

SEFSC Trawling Surveys

- Gulf of Mexico SEAMAP Trawling Surveys
 - Shrimp/Bottomfish
 - Summer and Fall
- Small Pelagics/Acoustics
- SEAMAP-South Atlantic Coastal Trawl Survey



NOAA
FISHERIES
SEFSC

Gulf of Mexico SEAMAP Trawling Surveys

- Shrimp/Bottomfish
 - Summer
 - Fall
- Small Pelagics/Acoustics
 - Fall

Objectives - Shrimp/Bottomfish

- Summer and Fall
 - Assess the abundance and distribution of demersal fauna in the US Gulf of Mexico
 - Collect size, sex, maturation and life history data of sampled species
 - Collect ichthyoplankton samples and data
 - Profile the water column at each sample site for temperature, salinity, dissolved oxygen, percent light transmission and fluorescence
- Summer Only
 - Estimate indices of relative abundance and population size structure of brown, white and pink shrimp off the Texas coast
 - Collect hydrographic data to map the hypoxic zone which occurs in the northern GOM

Summary - Shrimp/Bottomfish - Summer

- Time series: 1982-present
- Collaborating agency: SEAMAP-Gulf of Mexico
- 40 days at sea - NOAA Ship *Oregon II*
- Area of focus: US Gulf of Mexico, Texas to Florida
- Depths: 9-110 meters (5-60 fathoms)

Summary - Shrimp/Bottomfish - Fall

- Time Series: 1972-present
- Collaborating agency: SEAMAP – Gulf of Mexico
- 41 days at sea – NOAA Ship *Oregon II*
- Area of focus: US Gulf of Mexico, Texas to Florida
- Depths: 9-110m (5-60 fms)

Summary - Small Pelagics/Acoustics

- Time Series: 2002-present
- 40 days at sea – NOAA Ships *Pisces* and *Gordon Gunter*
- Area of focus: US Gulf of Mexico, Texas to Florida
- Depths: 50-500m (27-273fms)

Data QA/QC –

Shrimp/Bottomfish and Small Pelagics/Acoustics

Gulf of Mexico SEAMAP Trawling Surveys

- Data managed by SEAMAP-GOM / NMFS
- Field Party Chiefs are responsible for all quality assurance/quality control procedures prior to submitting the data to the IT Unit
- The database structure uses relational structures between tables and primary key and validation rules within tables in order to maintain QA/QC.
- Data are submitted to the Unit Lead of the IT Unit
- The IT Unit ensures that data are formatted correctly before submitting to the Data Manager, GSMFC.

SEAMAP-South Atlantic Coastal Trawl Survey

Multi-species survey targeting species utilizing unstructured habitats in nearshore waters



NOAA
FISHERIES



Summary Information

- Time series: 1986-present
- Coordinating agency: SEAMAP-South Atlantic
- Area of focus: Cape Hatteras, NC to Cape Canaveral, FL
- Depths: ~ 4.5-9m (15-30')



Data QA/QC

Data managed by SEAMAP-SA / SCDNR

- Each program is responsible for all quality assurance/quality control procedures prior to submitting the data in the correct format for the SEAMAP database.
- The database structure uses relational structures between tables and primary key and validation rules within tables in order to maintain QA/QC.
- Data are collected from data providers by the designated Data Coordinator.
- The Data Coordinator ensures that data are formatted correctly before forwarding those data to the Data Manager for uploading to the Oracle database.

SEFSC Plankton Surveys

Survey	Description	Area of Operation	Frequency
SEAMAP Plankton Survey	Surveys the occurrence, abundance and geographical distribution of the early life stages of fishes. The pelagic habitat of fish larvae are recorded through measurements of various physical and biological parameters. Map the distribution of fish eggs along the cruise track using a Continuous Underway Fish Egg Sampler (CUFES).	Coastal, shelf and open ocean waters of the Gulf of Mexico	Annual, Feb-Mar (Winter), Apr-May (Spring), and Aug-Sep (Fall)
SEAMAP GOM Offshore Ichthyoplankton survey (Mississippi and Louisiana)	Surveys ichthyoplankton eggs and larvae in the spring and fall (May and September, respectively) off Alabama, Mississippi, and Louisiana.	State and federal waters off Alabama, Mississippi, and Louisiana.	Annual, May-Jun and Sept-Oct
SEAMAP-GOM Ichthyoplankton survey (Alabama & Florida)	Ichthyoplankton surveys to collect larvae for red drum, king mackerel and other species. Florida has not collected since 2001.	State waters off Alabama out to 360 ft in depth. There are nine fixed stations near Mobile Bay, AL and from Alabama to Dry Tortugas, FL.	Annual, Aug-Sept

SEFSC Plankton Surveys



Objectives:

1. Assemble and maintain a consistent time series of data/observations on the occurrence, abundance, distribution and habitat of the early life stages of fishes, decapod crustaceans and invertebrate zooplankton in the Gulf of Mexico
2. Collect environmental data in order to determine the influence of those factors on distribution patterns, survival and, therefore, estimates of abundance of select taxa
3. Provide annual estimates of larval fish abundance from SEAMAP surveys for use as fishery independent indices of spawning stock size



Summary

- Time series: 1982 to present
- Current DAS/Vessels: ~30-35days for each of three dedicated plankton surveys on NOAA vessels; 2-5 days per state partner on Gulf state vessels
- Area: NMFS Gulf-wide, shelf and/or offshore waters out to U.S. EEZ (depending on survey); Gulf state agencies, state coastal waters primarily

Data Collection and QA/QC

At sea:

- pertinent station/environmental/sample collection data recorded on field data sheets until 1999
- now both manual and direct sensor input are recorded via SCS on 'electronic data sheets' or events (e.g. bongo event, CTD event, CUFES event etc.)
- direct ingestion of data after conclusion of an event into an Access database since 2002
- manual entry and editing by FPC of this database
- samples are preserved and sample numbers assigned



Data QA/QC

Mississippi Laboratories

- station/sample data verified & edited by Data Management group in consultation with FPC
- samples sent to National Marine Fisheries Research Institute of Gdynia and the Plankton Sorting and Identification Center in Szczecin, Poland (NMFRI/ZSIOP) for processing (two shipments per year)
- station/sample data sent to SEAMAP Archiving Center (SAC) at Florida Fish and Wildlife Research Institute (FWRI) for specimen data entry
- merged specimen/station/sample data error checked, verified, edited and data-based
- identification of select taxa are verified by SEFSC ichthyoplankton specialists



Data flow and QA/QC

FWRI/SAC in St. Petersburg, FL

- receives identified specimens from Poland
- specimens accessioned; id, count and length data entered into database
- these data files sent to MS Labs where it is merged with station and sample data
- specimens are loaned for examination by ichthyoplankton specialists

Data flow and QA/QC

At the National Marine Fisheries Research Institute of Gdynia and the Plankton Sorting and Identification Center in Szczecin, Poland (NMFRI/ZSIOP)

- samples sorted for fish eggs and larvae, decapod crustacean larvae and select invertebrate zooplankton
- fish larvae id'ed (most to family), counted, measured, and sample displacement volume measured
- specimens and data sheets sent to FWRI/SAC
- tuna larvae and data sheets airmailed to Miami Laboratory
- data files emailed to Pascagoula Laboratory

Data and Specimen QA/QC

- identification of select taxa (tunas, snappers, mackerels) are verified by SEFSC ichthyoplankton specialists before use of data in stock assessments



NOAA Vessel

Plankton samples collected; data retrieval and entry via SCS (Events); data ingested into Access database; manual entry and editing by FPC

SEAMAP sample nos. are assigned and keyed into database at sea on dedicated plankton surveys (or at the Lab for trawl surveys)

Sample inventory prepared at sea or at the Lab as noted above

Mississippi Labs Pascagoula

Electronic copy of the data are received after the cruise

Samples are received and keyed in shipment database (as required). Crates and reports (sample no., gear, and destination) are prepared.

Samples and spreadsheet are shipped to Polish Sorting Center

Spreadsheets (via email) and copies of original ID data sheets (via mail) received from Poland and stored. Invertebrate specimens held temporarily before shipment to SIPAC

Cruise/collection data are sent to SAC for entry of ID data.

Isa/If data are merged with cruise & collection data; error checks are made. Completed cruise data are data based.

If taxonomic updates are made, electronic copy of data and paper update sheets are sent to SAC.

If changes are made to collection or isa/If data after analysis and/or re-examination by taxonomic experts at MS Labs, cruise data are reprocessed and data based again.

SEAMAP Archives St. Petersburg, FL Ocean Springs, MS (SAC and SIPAC, SEAMAP Invertebrate Archiving Center)

Select samples and identified invertebrate zooplankton specimens sent to SIPAC for archival.

SAC receives corresponding cruise/collection data from MS Labs and enters specimen identifications, counts and length (isa/If data) into database.

Specimens are accessioned and made available for study; isa/If data are forwarded to MS Labs.

If changes are made to isa/If data files after re-examination by taxonomic experts other than at MS Labs. Isa/If files are revised at the Archive. These files are made available to MS Labs at FWRI (SAC) ftp site.

National Marine Fisheries Research Institute Plankton Sorting and Identification Center Gdynia & Szczecin Poland

Plankton samples sorted for fish eggs & larvae and invertebrate zooplankton (from select cruises). Specimens and data (IDs, counts, and lengths) are shipped to SAC. Invertebrate specimens & data are shipped to MS Labs, Pascagoula. Progress reports of sample processing are sent to the US quarterly.

'Chain of custody' for SEAMAP plankton samples, data and specimens taken during resource surveys



NOAA FISHERIES



NOAA
FISHERIES
SEFSC

Improving The NOAA NMFS and ICCAT Atlantic Bluefin Tuna Fisheries Management Decision Support.

Years: 2008-2011, Spring

Funding source: NASA

Geographical Area: Gulf of Mexico, Western Caribbean

Objectives: Develop habitat models for larval tuna to predict areas of abundance and improve the larval index for Atlantic bluefin tuna.

Management And Conservation Of Atlantic Bluefin Tuna (*Thunnus Thynnus*) And Other Highly Migratory Fish In The Gulf Of Mexico Under IPCC Climate Change Scenarios: A Study Using Regional Climate And Habitat Models

Years: 2012-2015, Spring

Funding Source: NASA

Geographic area: Gulf of Mexico, Caribbean.

Objectives: Enhance the management of multiple important highly migratory pelagic fish species in the Gulf of Mexico and surrounding waters, with particular focus on Atlantic bluefin tuna and other highly migratory tunas and billfishes.

QA/QC

- Larval fish are identified at the Miami Early Life History Laboratory, and at the Polish Sorting Center. All bluefin tuna identifications are checked. Questionable specimens are sent to the University of South Carolina for genetic identification
- Specimens are stored at the SEAMAP Archival center and at the Miami Laboratory

SEFSC Longline Surveys

- Fisheries Independent Bottom Longline Surveys
- Fisheries Independent Pelagic Longline Pilot Study Surveys
- Congressional Supplemental Sampling Program



**NOAA
FISHERIES**

NMFS

SEFSC/MS

Labs

Fisheries Independent Bottom Longline Surveys

1995 – 2012; 32 Surveys

U.S. GOM and Atlantic Coast

(Mexico and Cuba GOM, Cuba and Navassa Island
Caribbean)

- **OBJECTIVES**

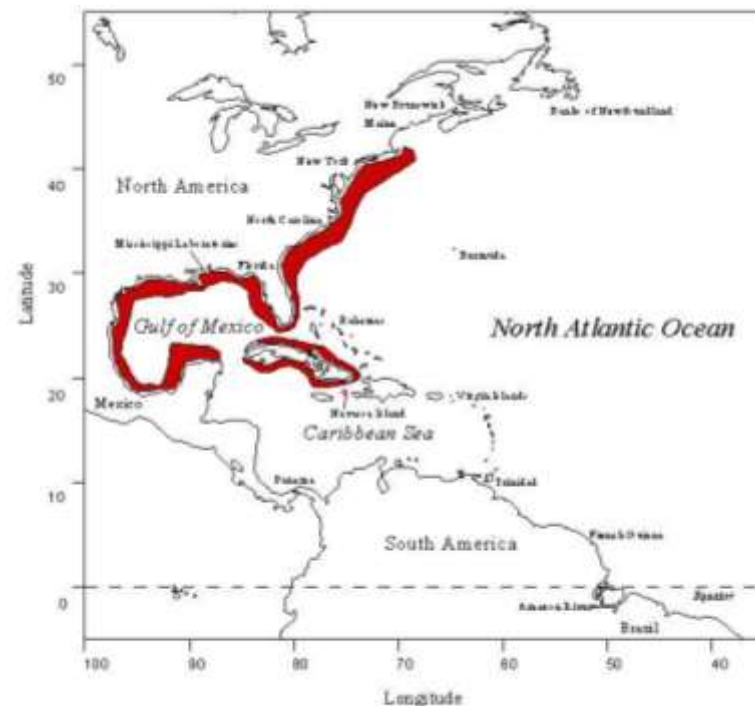
- Relative abundance and interannual variability
- Control potential sources of survey bias
- Repeatable survey
- Standardized survey design and gear
- Random station selection
- Biological sampling and tagging
- Associated environmental data
- 1999 objectives expanded to include red snapper

Summary

MS LABS CONTRIBUTION

TIME SERIES	# SURVEYS	# SETS	# CAPTURES	TAGS	FIN CLIPS	OTOLITHS	GONADS	# TAXA
1995 - 2012	32	4,232	38,340	7,575	5,969	2,328	2,075	152

- 60 Days At Sea Allocation
 - July, August, September
- NOAA Ships
- Survey Depth Range
 - 9 – 183 m Atlantic Ocean (primarily south of Cape Hatteras)
 - 9 – 366 m Gulf of Mexico (1997, 1998 Mexico, 1998 Cuba)
- Survey Design (proportional allocation of stations based on continental shelf width for 1° of latitude or longitude)
 - Atlantic; 9 – 55 m 60%, 55 – 183 m 40%
 - GOM; 9 – 55 m 50%:, 55 – 183 m 40%, 183 – 366 m 10%





NOAA
FISHERIES
SEFSC

Fisheries Independent Pelagic Longline Pilot Study Surveys

2004, 2005, and 2006

U.S. Eastern GOM and Atlantic Coast



- Pelagic longline gear
- Random station selection
- Repeatable survey
- Control potential sources of bias
- Species distributions
- Biological sampling and tagging
- Associated environmental data

Summary

PELAGIC LONGLINE PILOT STUDY SURVEYS

- U.S. Atlantic Coast (2004 and 2006)
- U.S. Gulf of Mexico (2005)
- 340 h effort
- 9316 hooks
- 92 sea days

CATCH SUMMARY

- 605 Captures
- 269 Tags Deployed
- 133 Fin Clips for DNA
- 22 Gonad
- 16 Otoliths





NOAA
FISHERIES

Congressional Supplemental Sampling Program

- FMP Species:
 - Increase sample size and precision of all estimates
 - Age-specific indices of abundance, growth, & fecundity
 - Gear & hook selectivity
- Habitats:
 - Map bottom topography using side scan sonar
 - Expand universe of reef habitats for annual reef fish surveys



Summary Information

Time series: 2011

Coordinating agency: SEAMAP-Gulf of Mexico

Area of focus: Continental shelf and shelf-edge banks

Dry Tortugas to US-Mexico Border

Depths: ~ 10-150m

Duration: 712 days at sea

Platforms: Chartered commercial fishing vessels



Longline Data QA/QC

Real-time data editing (at-sea)

- Specialized SELLIT data entry software (SCS/FSCS component)
- On-deck touch-screen data collection for all set and haul activities
 - Gear deployment and retrieval
 - Hook status
 - Preliminary catch data error checks
 - Error checking defaults for set and haul events (sequence of events)

Post-event data editing (at-sea)

- ACCESS supported data editing routines
 - Event data
 - Catch data
 - Environmental data
 - Photographic documentation

Laboratory-based data editing

- Reconciling catch and sampling-related data
- Respond to data analysis error checks
- Maintain the primary data archive
- Respond to data requests

SEFSC Video, Trap, and Vertical Line Surveys

- SEAMAP-Gulf of Mexico Reef Fish Video Survey
- MPA Reef Fish Video Survey
- Panama City Lab Fishery-Independent Reef Fish Survey
- South Atlantic Reef Fish Trap/Video Survey



NOAA
FISHERIES

SEAMAP-Gulf of Mexico Reef Fish Video Survey

Objectives

- Survey reef fish species associated with hard bottom, high relief, topographic features of the Gulf of Mexico
- Provide indices of relative abundance for use in stock assessments



Summary Information

Time series: 1992-1997, 2001-2002, and 2004-present

Coordinating agency: SEAMAP-Gulf of Mexico

Area of focus: Continental shelf and shelf-edge banks

Dry Tortugas to US-Mexico Border

Depths: ~ 40-150m

Duration: 60 days at sea

Platforms: NOAA ships and NMFS small boats





NOAA
FISHERIES

MPA Reef Fish Video Survey

Objectives

- Survey reef fish species associated with Madison-Swanson and Steamboat Lumps Marine Protected Areas
- Provide information on spawning aggregations of gag grouper (*Mycteroperca microlepis*) within MPAs



Summary Information

Time series: 2001 - 2010, 2012 - present

Coordinating agency: Gulf of Mexico Fishery Management Council

Area of focus: Marine protected areas in the eastern Gulf of Mexico

Region: Madison-Swanson and Steamboat Lumps MPAs

Depths: ~ 50 - 150m

Season: winter (Feb-April) to capture grouper spawning activity

Duration: 10-20 days at sea, weather critical

Platforms: NMFS small boats



Data flow and QA/QC

Shipboard collections

- SCS site, event, and environmental information
- Manual entry of biological collections (trap and vertical line) into access data base
- Initial QA/QC of shipboard collections by IT
- Secondary check by FPC.

Laboratory processing

- Laboratory read of video abundance data (mincount)
- Video measurement system (VMS) fish length measurement
- Manual entry of mincount and length data into access data base
- Review includes secondary read of mincount and length data from video
- QA/QC of entered data by primary data base manager
- Distribution to analysts for inclusion in primary in data set

Data flow and QA/QC

Archiving

- Primary physical videos and station sheets stored in secure environment
- No current backup
- Hi-8, mini-DV, and Blu-ray formats
- File sizes are extremely large and backup is expensive
- Event, environmental, and biological data backed up at MS Laboratories
- All historic relational data tables converted to and stored in Access
- Analysis sets built in SAS and converted for use in GIS applications
- Analysis builds handled as yearly append to primary analysis set

Issues

- Disjunction between storage and analysis sets
- Historically the SCS event data stored separately from video read data
- Updates to analysis sets are often time sensitive and piece-meal
- Poor feedback to primary tables

Future

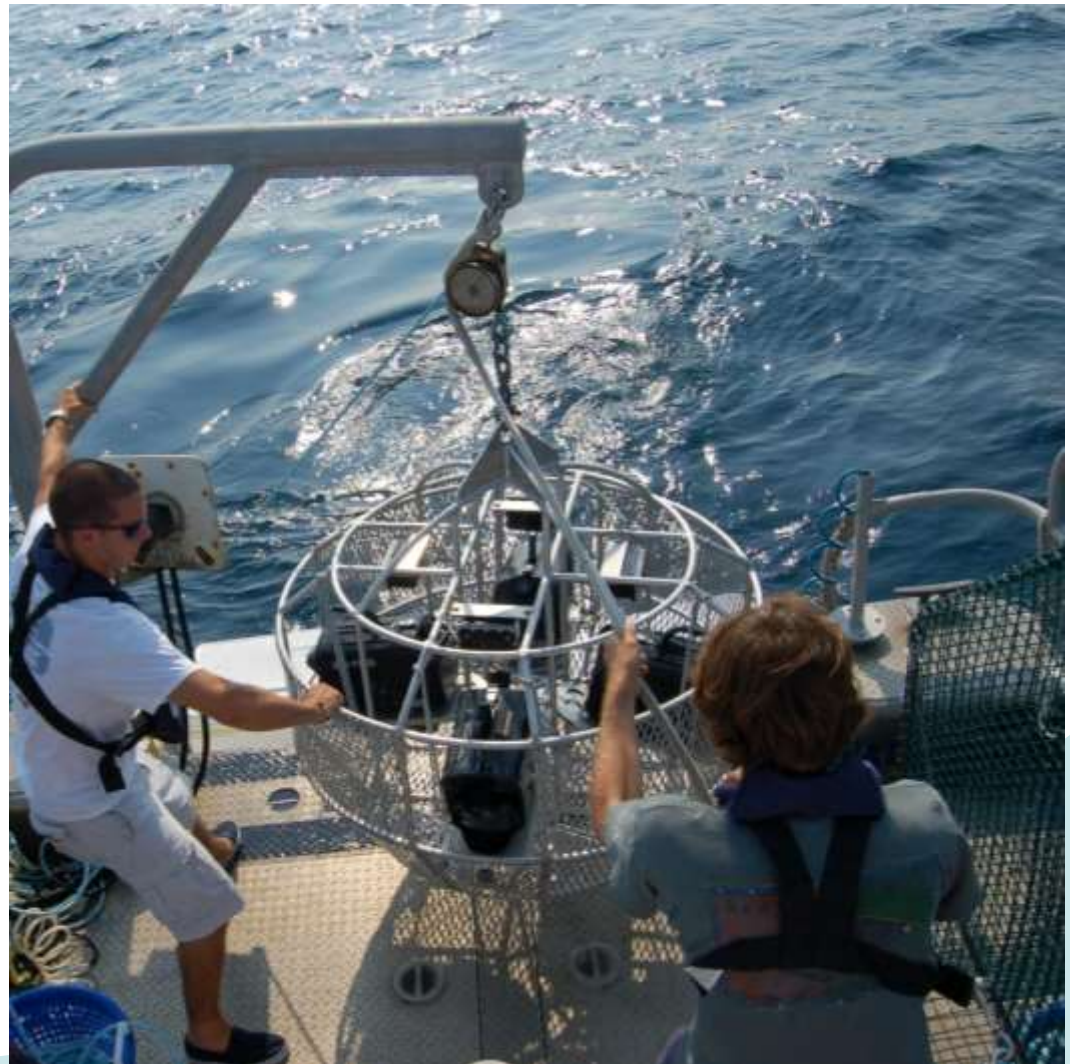
- Mainframe storage of historic physical videos as a true backup (Stennis)
- Centrally located and controlled 'gold standard' distribution set (MS Labs)
- Oracle
- Analysis builds and distribution directly from primary 'gold standard' set



NOAA
FISHERIES

Panama City Lab Fishery-Independent Reef Fish Survey

Summary: Survey the abundance of reef fishes occurring on rocky reefs and live bottom in inner and mid-shelf waters (8-50m) off the Florida Panhandle and Big Bend Region.



Data Flow and QA/QC

- Tape with best view of the reef is analyzed, and if all are equally good, then one is randomly selected.
- First unoccluded 20 min of the 30 min of tape are analyzed.
- Fish id'd to lowest discernible taxon.
- Estimator of abundance is the maximum number of a given species in the field of view (= min count) at any one time during the 20 min. being reviewed.
- Fish lengths are obtained from stereo images using VMS software, and only from a frame showing the min count to eliminate the possibility of double measuring.
- The first 10 tapes analyzed each year are read by both readers to ensure consistency & identify issues in species id, min counts, and habitat classification.
- Each year 10 tapes from each lab involved in the West Florida shelf cooperative survey are also read by the other 2 labs, again to ensure consistency & identify issues in species id, min counts, and habitat classification.
- All data are proofed and entered in an Access database onsite.
- Tapes with no evidence of hard bottom nearby, or with view obscured by poor visibility or bad camera angle, are excluded from CPUE analyses.

A background image showing several fishing boats docked at a pier. The boats are white with blue and yellow accents. One boat has the number '23087' visible. The scene is set in a harbor with a blue sky and water.

South Atlantic Reef Fish Trap/Video Survey



NOAA
FISHERIES

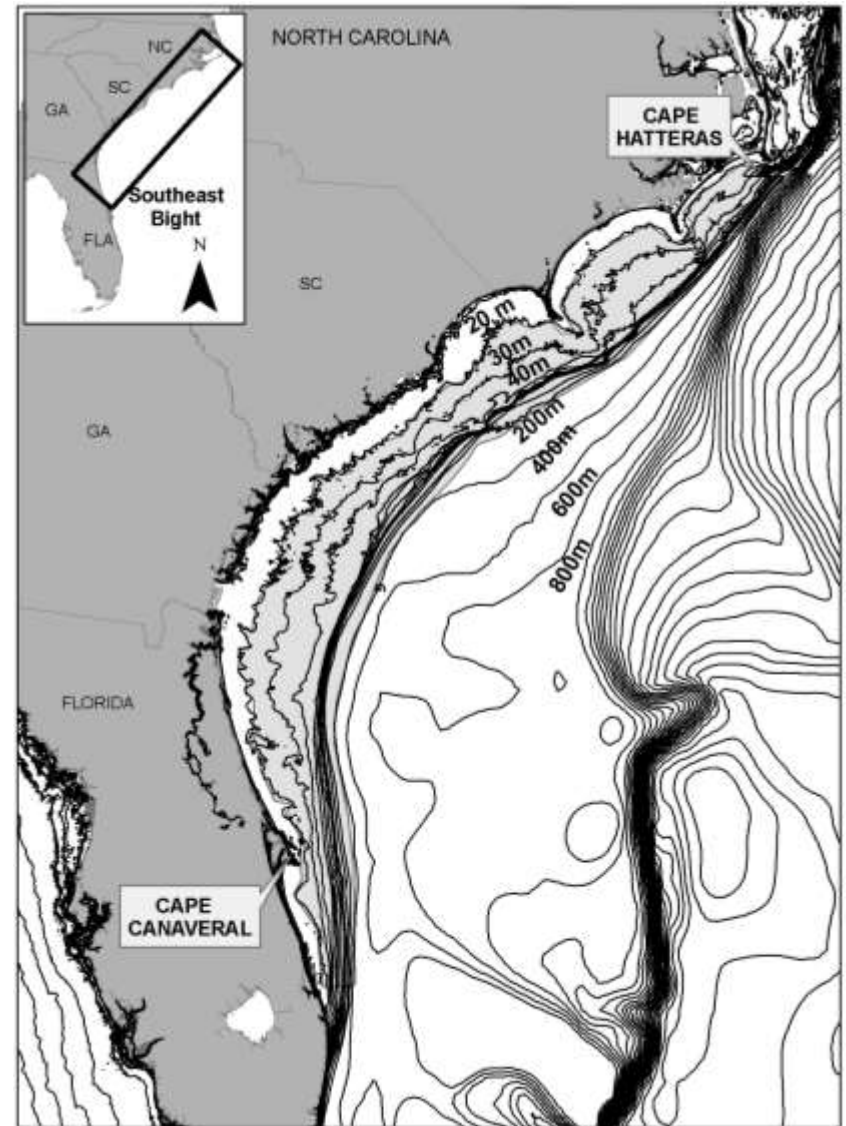
Multi-species survey targeting species within
the snapper-grouper complex (> 100 species)

Focal species are hardbottom-associated; thus,
survey targets hardbottom habitats



Summary

- Time series:
 - Chevron trap survey: 1990-present
 - Video survey: 2010-present
- Coordinating programs/agencies:
 - MARMAP/SEAMAP-SA (SCDNR; 1972-present)
 - SEFIS (NMFS SEFSC; 2010-present)
- Both groups work cooperatively to perform
- Trap-video survey in SA waters
- Area of focus: Cape Hatteras, NC to Port St. Lucie, FL
- Depths: ~ 15 to 100m





NOAA
FISHERIES

Data QA/QC

Data managed by both MARMAP (SCDNR) and SEFIS (SEFSC); agency-specific QA/QC procedures

MARMAP/SEAMAP-SA

All station, length-frequency, age, reproductive information, video entry data are reviewed at various levels for accuracy by

1. Data entry staff and Chief Scientist
2. Database managers
3. Age and reproduction scientists and species leaders
4. Program coordinator



NOAA
FISHERIES

Data QA/QC

Data managed by both MARMAP (SCDNR) and SEFIS (SEFSC); agency-specific QA/QC procedures

SEFIS

All station, length-frequency, age-growth, video log, video entry data are reviewed three times for accuracy (by 1. data entry person, 2. database manager/chief scientist, and 3. SEFIS coordinator)

For each reader, all videos are re-read by video/fish ID expert until no systematic errors occur, after which time 10% of videos are re-read by video/fish ID expert

Caribbean Surveys, Visual Surveys, and Other Surveys of the SEFSC

- Visual Coral Reef Habitat and Reef Fish Surveys - Florida Keys, Dry Tortugas, and Southeast Florida Shelf
- Visual Coral Reef Habitat and Reef Fish Surveys - U.S. Caribbean
- SEAMAP Queen Conch Visual Survey - U.S. Caribbean
- Caribbean Reef Fish Video Survey
- Gulf of Mexico Shark Pupping and Nursery (GULFSPAN) Survey



NOAA
FISHERIES
SEFSC

Visual Coral Reef Habitat and Reef Fish Surveys - Florida Keys, Dry Tortugas, and Southeast Florida Shelf

Objective: These surveys collect data concerning abundance and size of reef fishes, and the characteristics of bottom habitat features in order to better understand how natural and manmade stressors are changing reef fish populations and communities.

Collaborations

This survey is a multi-agency effort to monitor reef fish populations in the Florida Keys and Dry Tortugas.

Agencies involved include:

- NOAA SEFSC
- Florida Fish and Wildlife Conservation Commission's Florida Fish and Wildlife Research Institute (FWRI)
- University of Miami's Rosenstiel School of Marine and Atmospheric Science (UM-RSMAS)
- National Park Service (NPS).

This collaborative effort is the culmination of nearly three decades of independent Florida Keys monitoring programs aimed at fish populations in the region.



Data Management and QA/QC

- Data collected for each stationary point count are entered via the RVC Data Entry Program following the methods outlined in SOP.
- Data for the Field/Boat log and Water Quality/Environmental log are kept on an agency-specific Excel spreadsheet.
- Initial data entry and proofing is the responsibility of the participants who collected the data in field.
- At the end of the field season, designated agency-specific data managers are responsible for assuring that all data collected have been entered and proofed.
- Each diver must maintain a minimum level of accuracy as described and defined in SOP.
- Data managers then distribute their agency's data to the NOAA and UM-RSMAS data managers, who will be responsible for data verification, storage, and distribution following steps in the SOP.



NOAA
FISHERIES
SEFSC

Visual Coral Reef Habitat and Reef Fish Surveys - U.S. Caribbean

These surveys collect data concerning abundance and size of reef fishes, and characterize bottom habitat features.

- Survey Design and Methods
 - U.S. Virgin Islands and Puerto Rico, <33m deep
 - Annual, May-Sep
 - Transect surveys (30 x 2m, modified AGRRA)
 - Stationary point count
 - 5 or 7.5 m radius
 - Bohnsack and Bannerot method
- Species Targeted
 - non-cryptic, diurnal reef fishes
 - macroinvertebrates
 - *Diadema* urchins
 - Spiny lobster
 - Queen conch

Data Management and QA/QC

- How Data Quality is Insured
 - training and observer cross comparisons
 - data entered, proofed, and checked vs datasheets
 - internal data checks (e.g., size limits)
- How the Data are Subsequently Managed
 - data are maintained by PI in lab database
 - PI manages analysis
- How are Data Used in Stock Assessments
 - trends in density, size distribution, “CPUE” – e.g., freq. of occurrence
 - changes in max or mean sizes observed



NOAA
FISHERIES
SEFSC

SEAMAP Queen Conch Visual Survey - U.S. Caribbean

These surveys collect data on queen conch size and abundance within the territorial sea of the U.S. Virgin Islands, Puerto Rico and the contiguous EEZ.

- Survey Methods
 - U.S. Virgin Islands and Puerto Rico territorial waters in 10-90 ft depths, some sampling occurs in federal waters.
 - Annual, USVI, Jun-Oct; PR, Jul-Nov
 - Queen conch abundance and density are estimated by visual census surveys conducted along predetermined compass headings by SCUBA divers using diver propulsion vehicles.
 - There is no extraction and/or collection of queen conch.
 - Radial surveys (10m radius)
 - Measure all conch (shell length and lip thickness)
- Species Targeted
 - Queen conch

Data Management and QA/QC

- How Data Quality is Insured
 - training and observer cross comparisons
 - data entered, proofed, and checked vs datasheets
 - internal data checks (e.g., size limits)
- How the Data are Subsequently Managed
 - data are maintained by PI in lab database
 - PI manages analysis
- How are Data Used in Stock Assessments
 - trends in density, size distribution, habitat-based extrapolation to standing stock, “CPUE” – e.g., freq. of occurrence, comparison with other f-I approaches.



NOAA
FISHERIES

Caribbean Reef Fish Video Survey

Objectives

- Survey reef fish species associated with hard bottom, high relief, topographic features of the Puerto Rico and the USVI
- Provide indices of relative abundance for use in stock assessments



Summary Information

Time series: 2009 & 2012

Coordinating agency: Southeast Fisheries Science Center

Area of focus: Continental shelf and shelf-edge banks

Region: Puerto Rico, St. Thomas, St Johns and St. Croix (USVI)

Depths: ~ 50 -150m

Duration: 30 days at sea

Platforms: NOAA ships



Data flow and QA/QC

Shipboard collections

- SCS site, event, and environmental information
- Manual entry of biological collections (trap and vertical line) into access data base
- Initial QA/QC of shipboard collections by IT
- Secondary check by FPC.

Laboratory processing

- Laboratory read of video abundance data (mincount)
- Video measurement system (VMS) fish length measurement
- Manual entry of mincount and length data into access data base
- Review includes secondary read of mincount and length data from video
- QA/QC of entered data by primary data base manager
- Distribution to analysts for inclusion in primary in data set

Data flow and QA/QC

Archiving

- Primary physical videos and station sheets stored in secure environment
- No current backup
- Hi-8, mini-DV, and Blu-ray formats
- File sizes are extremely large and backup is expensive
- Event, environmental, and biological data backed up at MS Laboratories
- All historic relational data tables converted to and stored in Access
- Analysis sets built in SAS and converted for use in GIS applications
- Analysis builds handled as yearly append to primary analysis set

Issues

- Disjunction between storage and analysis sets
- Historically the SCS event data stored separately from video read data
- Updates to analysis sets are often time sensitive and piece-meal
- Poor feedback to primary tables

Future

- Mainframe storage of historic physical videos as a true backup (Stennis)
- Centrally located and controlled 'gold standard' distribution set (MS Labs)
- Oracle
- Analysis builds and distribution directly from primary 'gold standard' set



NOAA
FISHERIES
SEFSC

Gulf of Mexico Shark Pupping and Nursery (GULFSPAN) Survey

Objective: Mid-water gillnet survey to monitor juvenile shark populations in the coastal Gulf of Mexico.

- **Summary**

- 2001 through current year
- April 1 through October 31
- 0700 – 1700

- **Gear**

- 600' (186 m) gillnet
- 6, 100' panels, 10' deep
- Variable stretched-mesh sizes (3.0 - 5.5")

- **Useful in Assessments**

- SAFE reports for NOAA's Sustainable Fisheries Division
- Yearly report to NOAA's Highly Migratory Species Office
- Catch series for shark stock assessments

- **Collaborations**

- NOAA Fisheries Panama City Laboratory (2001-current)
- GCRL at the University of Southern Mississippi (2003-current)
- Louisiana State University (2003-2005)
- FMNH at the University of Florida (2006-2011)
- Dauphin Island Sea Lab (2007-current)
- Florida State University CML (2007-current)
- University of West Florida (starting in 2013)

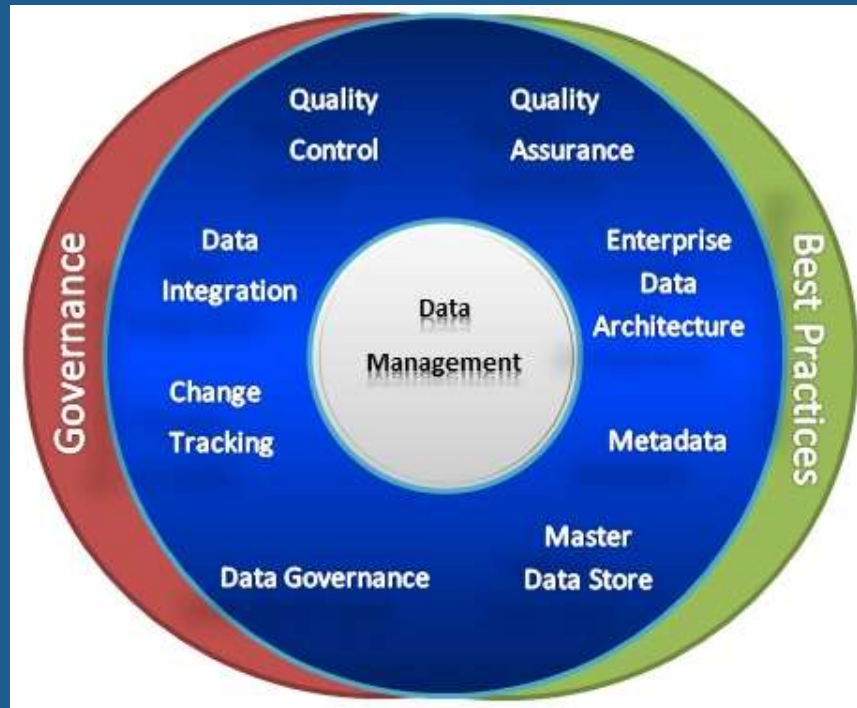


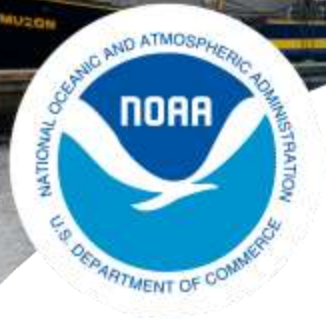
Database & QA/QC

- Collaborating institutions house their own data, run their own analyses (CPUE and habitat associations), and build their own tables and figures separately for end of year report.
- NOAA Fisheries Panama City Lab
 - Access Database
 - Co-PI enters data after each field expedition
 - Data proofed at the end of field season (Nov 1)
- NOAA Fisheries Panama City Lab organizes written section, tables, and figures of end of year report and then sends to HMS, usually in January.

SEFSC

Fisheries Independent Data Management





**NOAA
FISHERIES**

Fisheries Independent Data Management: Retrospect

Retrospect

Fisheries and Environmental Data Collection

- Over 50 years of fisheries independent data
- Standardized random sampling since 1972
- SEAMAP sampling protocols since 1982
- Used for stock assessment, habitat and ecosystem studies

Retrospect

Fisheries and Environmental Data Collection

- Solutions used for capturing, recording, converting and storing information varied throughout the years.
 - Field Data Sheets 1956 - 1987
 - SEAMAP DBMS V3 (ASCII Main Frame) 1988 - 1998
 - SEAMAP DBMS V4 (Oracle) 1998 to Present
 - Shipboard Computer System (SCS) 2000 – Present
 - Fisheries Scientific Computer System (FSCS 1.6) 2003 – Present
 - SEAMAP Data Entry System (SDES) 2003 - Present
- New system architectures introduced challenges for converting and storing data (e.g. data loss, data inexactitude, relational impedance of data sets)

Retrospect

Fisheries and Environmental Data Collection

- Station and catch data from 1956 to 1987 were recorded using an 80 column pseudo (punch card) format
 - Data sheets from 1956 thru 1967 converted to microfiche.
 - Data sheets from 1967 to present day are maintained by SEFSC.
 - All known punch cards have been converted to ASCII files.
- Fields sheets from 1967 to September 2005 were damaged by Katrina
 - Most have been restored
 - Some have been scanned
- The majority of SEFSC trawl data from 1972 to 1987 and all SEAMAP data since 1982 have been converted to an Oracle Database.
- Since 2003 the number of manual data sheets has been reduced through electronic data collection platforms.



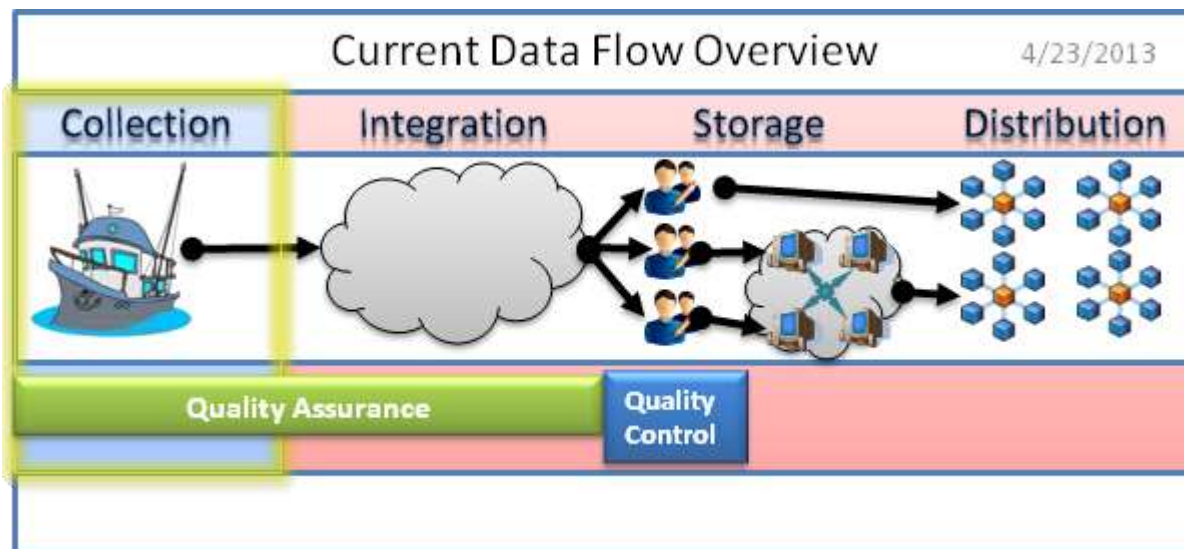
NOAA
FISHERIES

Fisheries Independent Data: Current Data Management Policy



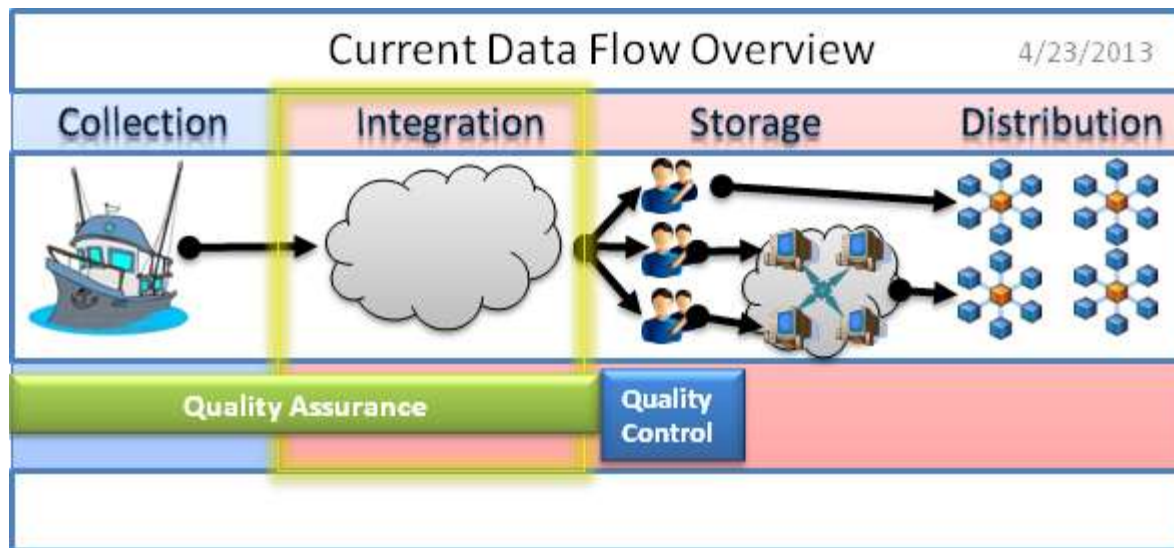
NOAA FISHERIES

SEFSC Information Management Practice



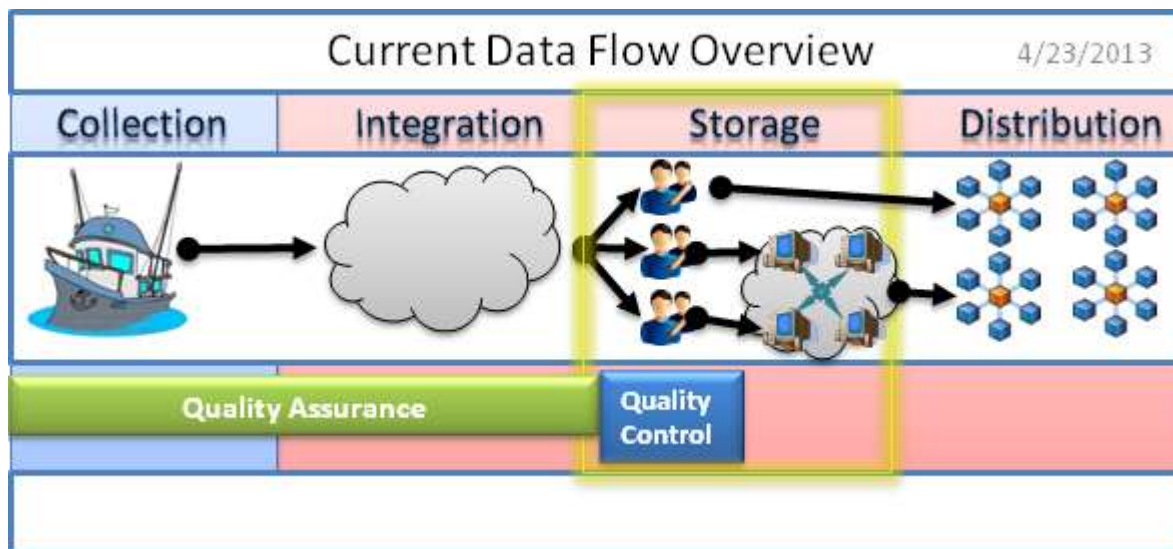
- Survey data are collected using separate application platforms
 - Shipboard Computer System (SCS) – environmental data
 - Fisheries Scientific Computer System (FSCS 1.6) – trawl data
 - SouthEast Longline Input Technology (SELLIT) – longline data
 - CTD event-based tracking system
 - SEAMAP Data Entry System (SDES)

SEFSC Information Management Practice



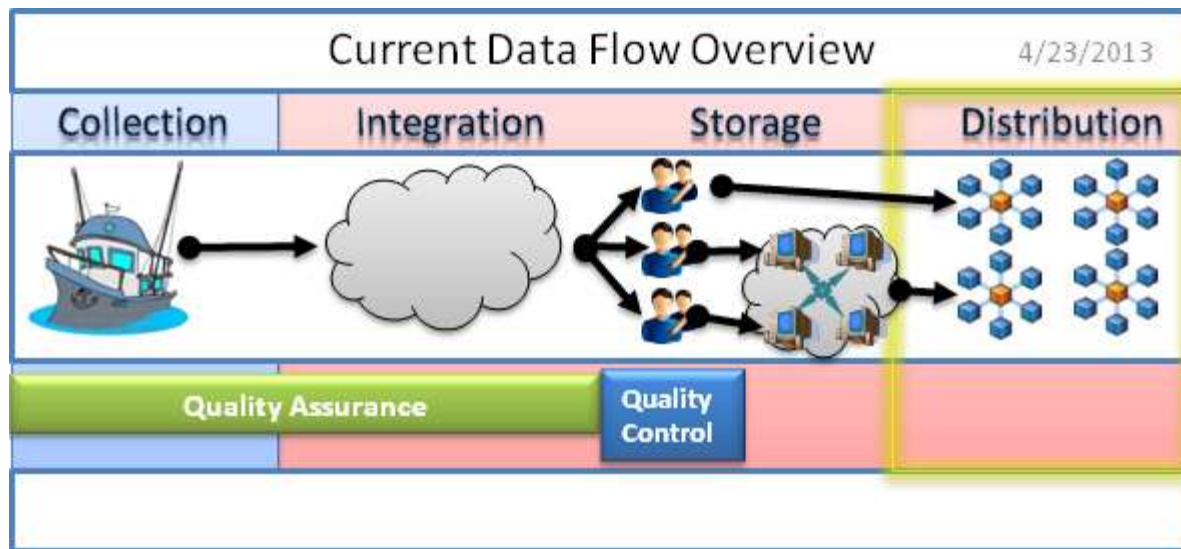
- data are collected from each of the data entry systems and consolidated for integration into a centralized/staging database.
 - SEAMAP Data Entry System (SDES) collects, consolidates and integrates data into on centralized database
- Quality assurance checks are applied during the collection and integration of data

SEFSC Information Management Practice



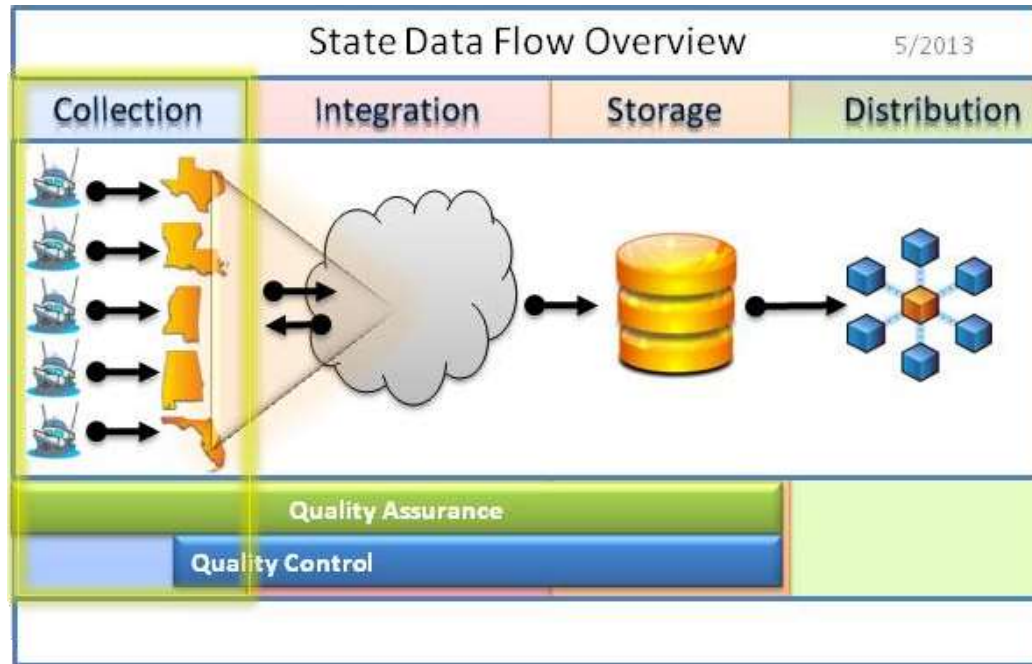
- Multiple copies of survey data are generated and distributed
- Separate Quality Control checks are applied to each distributed copy and, in some cases, changes do not recede to the originating source.
 - End result is the generation of multiple disparate data stores
- Specific scientific applications developed to supplement data copies.

SEFSC Information Management Practice



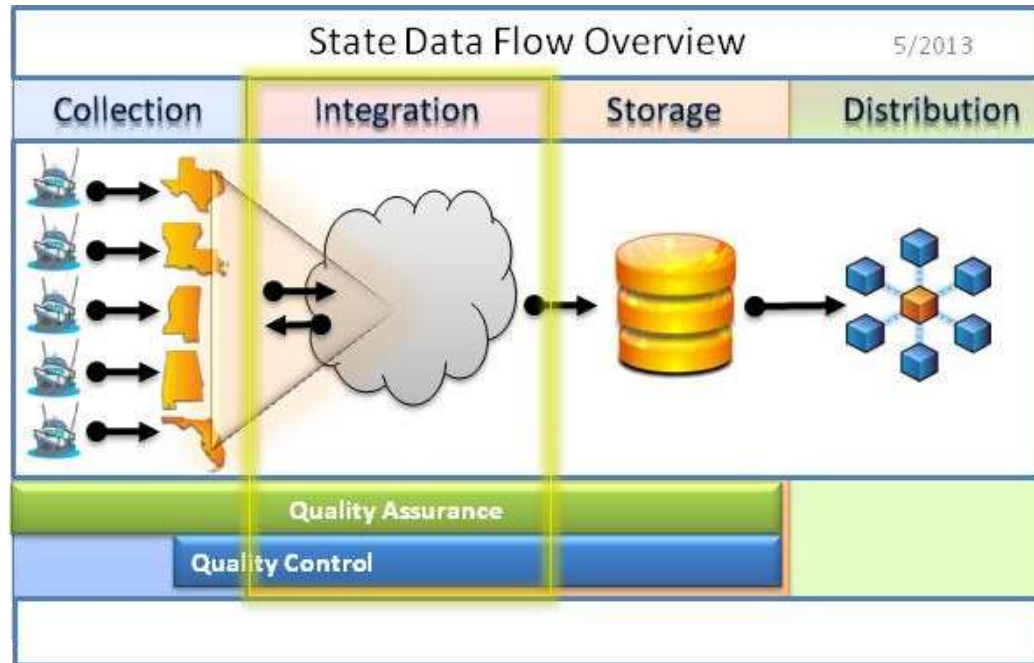
- No single source for distributing / sharing information.
- Data request must be channeled to individuals who are believed to have the “gold” copy.
- Inconsistencies can exist in information shared with / distributed to internal and external stakeholders

State Fisheries Independent Information Management Practice



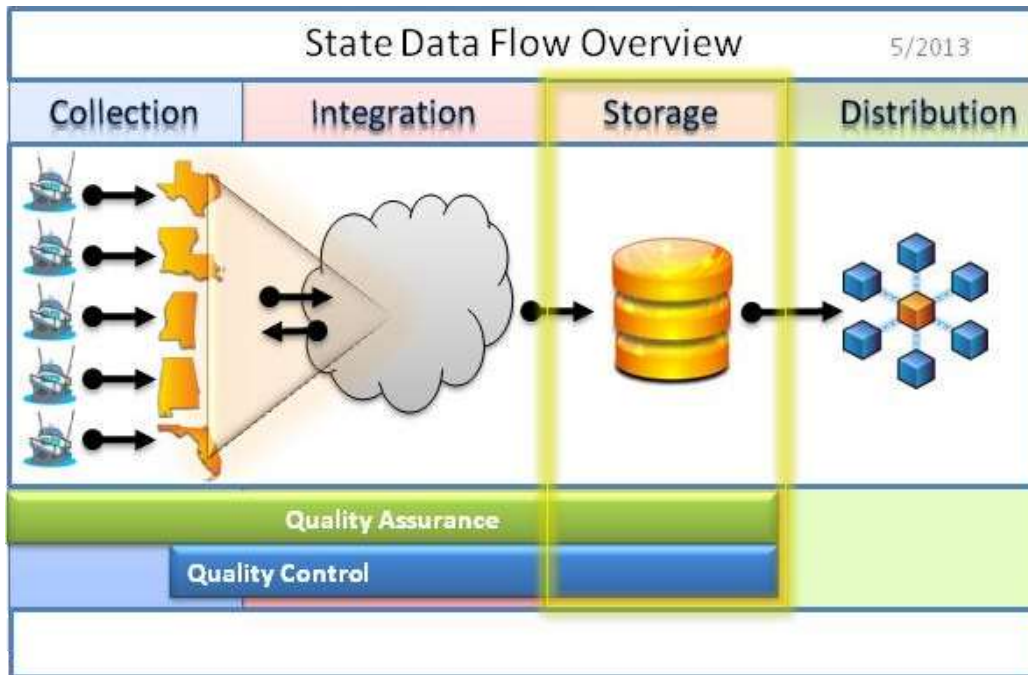
- State survey data are collected using several application platforms
 - Alabama, Mississippi, and Florida use the same, or similar, methods and software as the SEFSC for compiling survey data.
 - Texas and Louisiana utilize locally developed collection methods to compile survey data.
 - Louisiana has recently shown interest in adopting SEFSC methods for compiling survey data.

GSMFC Information Management Practice



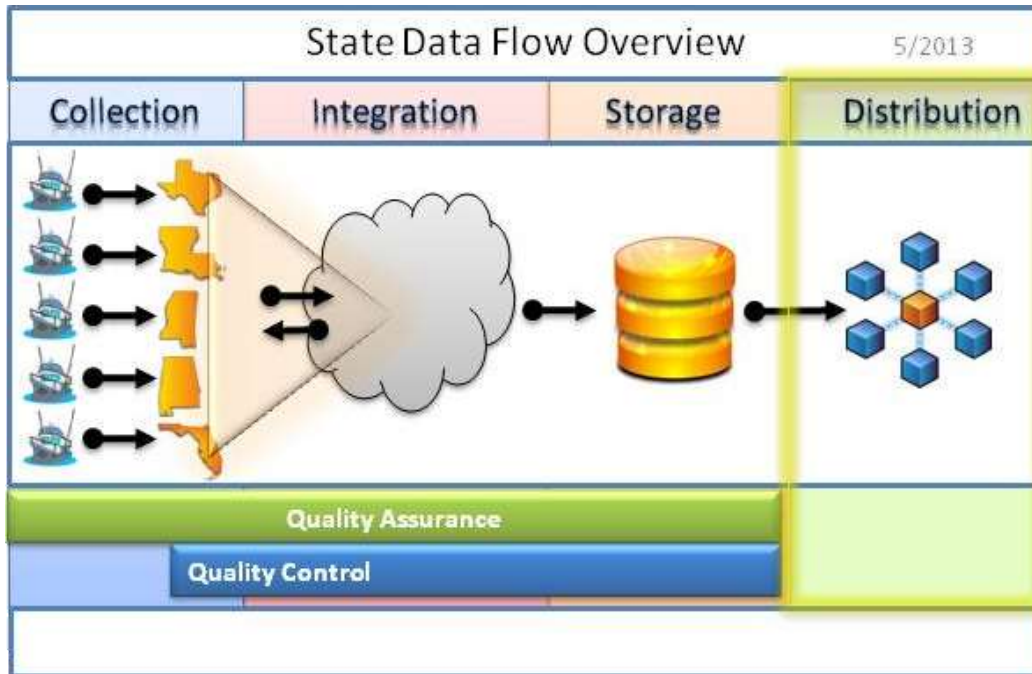
- State partners provide survey data to GSMFC for integration into their Master Data Store
 - GSMFC retains the responsibility for integrating raw data
 - data are validated by both the state partners and GSMFC
 - Ambiguous data identified by GSMFC is provided to, and approved by, state partners prior to implementation.

GSMFC Information Management Practice



- All Survey data collected by state partners and SEFSC SEAMAP data are integrated into a Master Data Store
 - While copies of the Master Data Store exist, all information is distributed from a single source
 - Ambiguous data discovered after integration is communicated back to state/federal partners for consideration

GSMFC Information Management Practice



- All distributed information is provided from a single source Master Data Store.
- Ambiguous data identified by the scientific community is verified by GSMFC and, if necessary, referred to state/federal partners for consideration.
- Changes / corrections to ambiguous data are redistributed upon request.



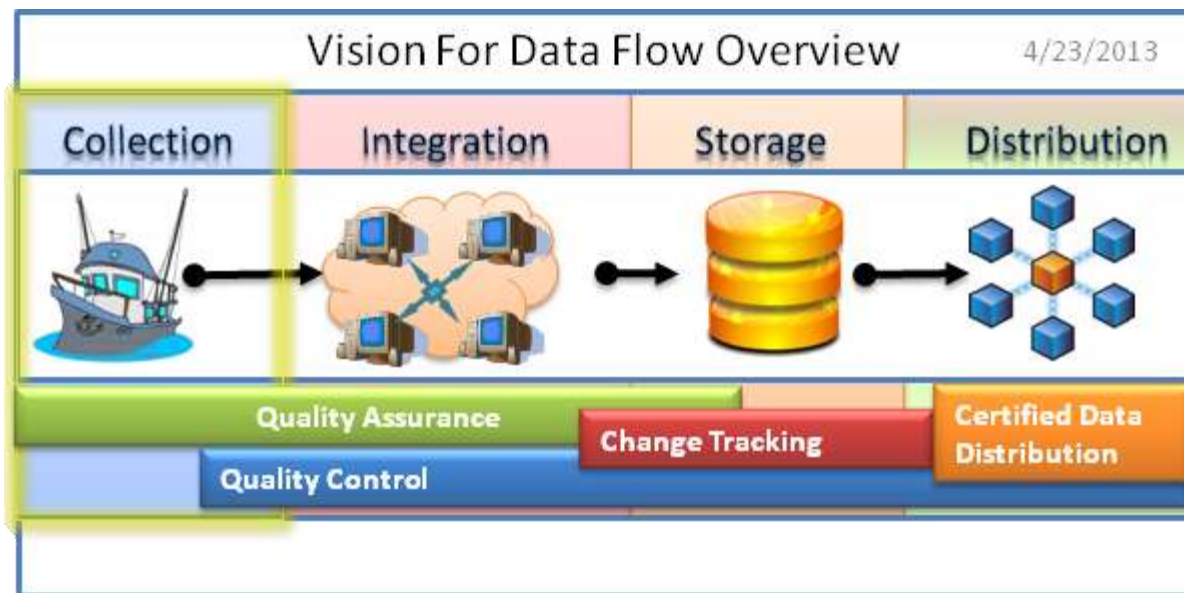
**NOAA
FISHERIES**

SEFSC Fisheries Independent Data Management Vision and Strategy



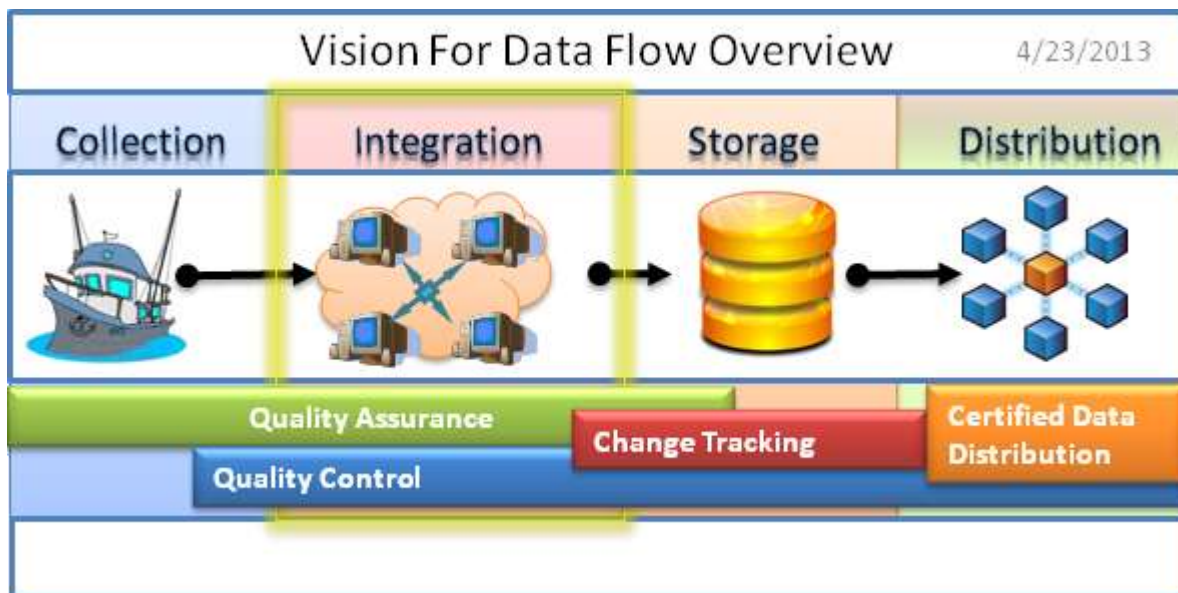
NOAA FISHERIES

SEFSC Vision and Strategy



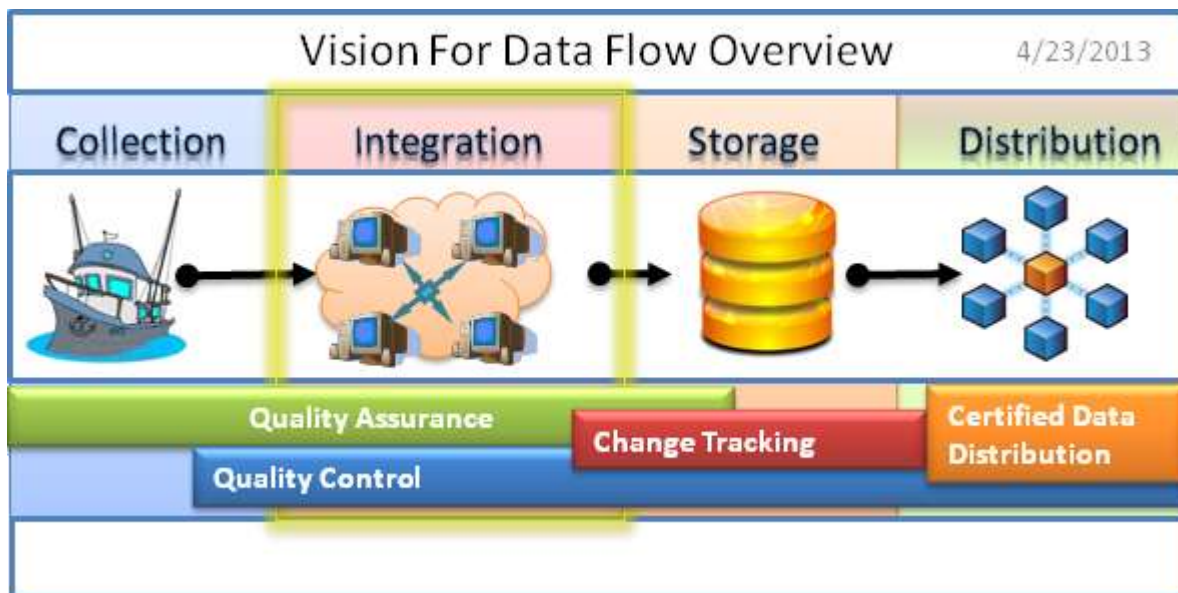
- Methods for collecting and extracting survey data will stay unchanged.
- SEFSC will collaborate with state and federal agencies to standardize collection and extraction methods for survey data.
 - Currently SEFSC is participating in the development of FSCS II
 - The SEFSC is working with NEFSC to assess / incorporate their software solutions for capturing Trawl survey data.

SEFSC Vision and Strategy



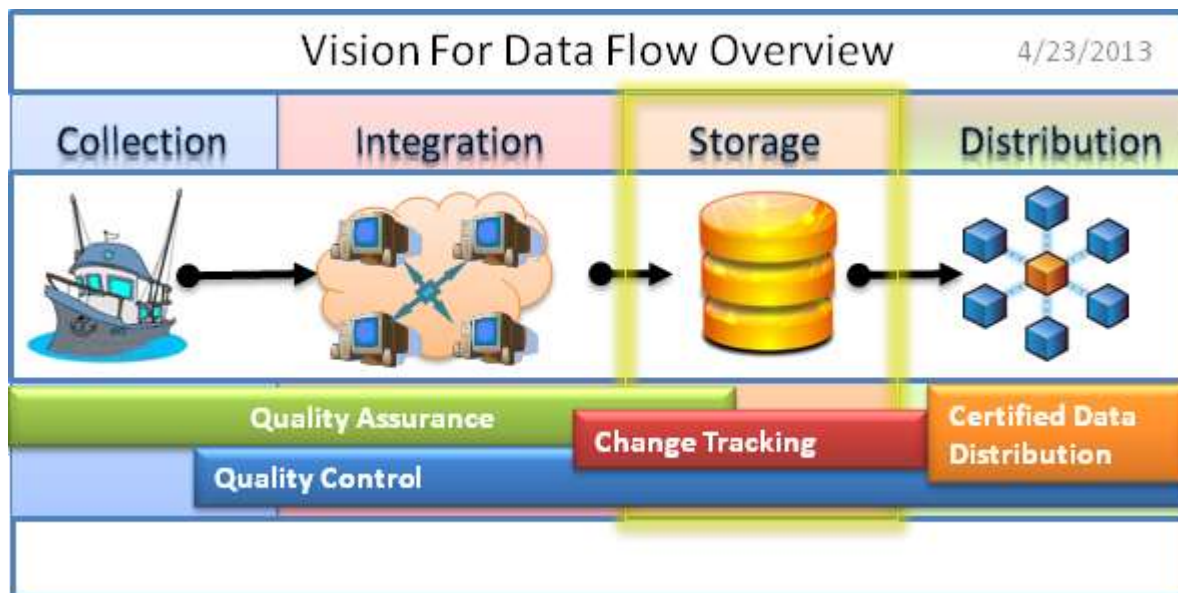
- Data will be staged prior to incorporation into the master data store
- Quality Assurance checks will be applied and measured from collection and continue until final incorporation into the master data store.
- Quality Control checks will be applied and measured from integration throughout the lifecycle of the survey data.

SEFSC Vision and Strategy



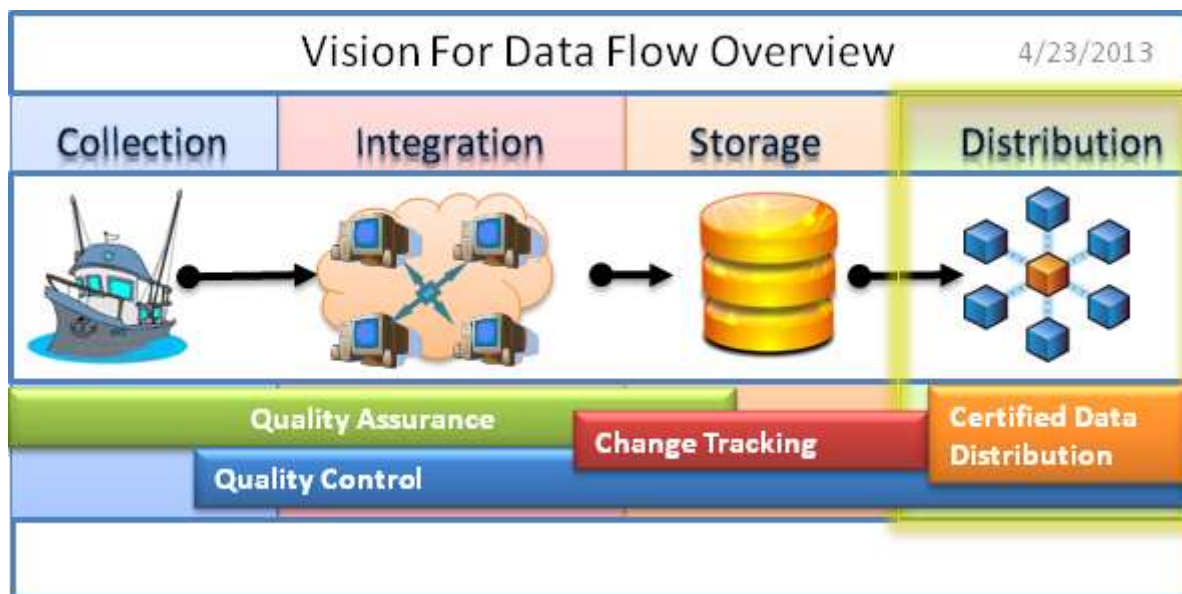
- Specific scientific applications will be developed by the SEFSC for adding supplementary data, and will be implemented as an element of data integration
 - Ensures that adopted quality assurance and quality control measures are applied prior to incorporation into the master data store

SEFSC Vision and Strategy



- Change Tracking will be implemented to protect the integrity and relevance of data contained in the master data store and to ensure that the interests of all stakeholders are protected.
 - All changes will be reviewed and approved by an appointed Change Control Board prior to implementation
 - Provides for the traceability and recession of data changes.

SEFSC Vision and Strategy



- All information will be distributed from the master data store.
- Ambiguous data identified by the scientific community will be referred to the Change Control Board for consideration.
- All changes as a result of Quality Control activity will recede back to the master data store for redistribution.

SEFSC Vision and Strategy

Software Development

- Incorporate agile software development methodologies that form a structured framework for planning and controlling the creation of survey specific information systems.
 - Software solutions evolve through collaboration between the scientific staff and IT personnel.
- Ensure that supporting documentation is written throughout the lifecycle of software development.
 - This documentation will serve as the basis for capturing and recording descriptive, structural and administrative metadata.
- Develop and share common software solutions through partnered relationships with other NMFS organizations.

SEFSC Vision and Strategy

Data Management

- Formalize a Data Management Plan depicting MSLabs governance over the collection, processing, administration, storage, safeguarding and distribution of information.
 - Impelled by the guidelines of FIMACs Fisheries Enterprise Data Management Vision and Strategy, dated March 13, 2013
 - Adherence to agency and fishery policy / procedural directives
 - *NOA 212-15: Management of Environmental Data and Information*
 - *NMFS Data and Information Management Policy Directive*
 - *Data Management Planning Procedural Directive*
 - *Data Documentation Procedural Directive*
 - *NMFS Data Management Plan Procedural Directive*
 - *NMFS Data Documentation Procedural Directive*

SEFSC Vision and Strategy

Partnerships

- SEFSC will advance information management effectiveness through collaborative relationships by leveraging the technological capabilities of other state and federal agencies.
 - SEFSC is currently working with the National Coastal Data Development Center (NCDDC).
 - Ensure that data are preserved and stewarded for the long-term.
 - Ensure the availability and accessibility of information to the scientific community.
 - Standardize data dissemination to improve the integration and/or interoperability of information with other analytical toolsets.
 - Ensure the scalability of IT infrastructures to reinforce information lifecycle management.

It is the goal of the SEFSC to establish strong, well-documented Quality Assurance and Quality Control measures to ensure and preserve the integrity, reliability and accessibility of our information and data.